

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 04-163497

(43)Date of publication of application : 09.06.1992

(51)Int.Cl.

G10L 3/00  
G10L 9/14  
// G10L 3/02

(21)Application number : 02-289113

(71)Applicant : RICOH CO LTD

(22)Date of filing : 26.10.1990

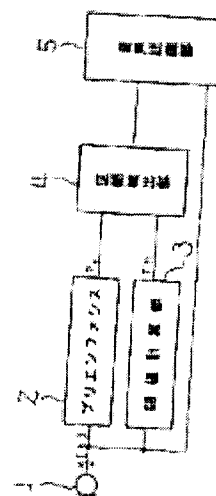
(72)Inventor : MATSUSHITA MITSUGI

## (54) VOICE SECTION DETECTING METHOD

### (57)Abstract:

**PURPOSE:** To reduce the occurrence of erroneous detection of a voice section without lowering a processing speed by detecting a section where a voice is present by means of an output through a preemphasis to emphasize the component of a frequency approximately equal to the characteristic frequency region frequency of a vowel having a low voice power.

**CONSTITUTION:** A voice input part 1 collects voices and converts the voices into an electric signal xt, and a preemphasis 2 emphasizes the component of a frequency approximately equal to the characteristic frequency region frequency of a voice to increase a voice of a vowel having a low voice power. A threshold calculating part 3 selectively calculates a threshold and calculates a threshold Th by means of an output signal Pr from the preemphasis 2 at a section where no voice is present. A section detecting part 4 detects a section by means of signals Pr and Tg, an output signal and an input signal therefrom are inputted to a voice input part 5 to recognize a voice.



⑩ 日本国特許庁(JP)

⑪ 特許出願公開

## ⑫ 公開特許公報(A) 平4-163497

⑬ Int. Cl.<sup>9</sup> 識別記号 庁内整理番号 ⑭ 公開 平成4年(1992)6月9日  
 G 10 L 3/00 3 0 1 A 8842-5H  
 9/14 D 8622-5H  
 // G 10 L 3/02 3 0 1 8842-5D  
 審査請求 未請求 請求項の数 1 (全4頁)

⑯ 発明の名称 音声区間検出方法

⑰ 特 願 平2-289113

⑱ 出 願 平2(1990)10月26日

⑲ 発 明 者 松 下 貢 東京都大田区中馬込1丁目3番6号 株式会社リコー内  
 ⑳ 出 願 人 株 式 会 社 リ コ ー 東京都大田区中馬込1丁目3番6号  
 ㉑ 代 理 人 弁 理 士 柏 木 明

## 明 細 書

1. 発明の名称 音声区間検出方法

## 2. 特許請求の範囲

音声を集音して電気信号に変換し、変換された信号を基に音声の存在する区間を検出するようにした音声区間検出方法において、音声パワーの小さい母音のホルマント周波数近傍の周波数成分を強調するプリエンファスを通した出力を用いて音声の存在する区間を検出するようにしたことを特徴とする音声区間検出方法。

## 3. 発明の詳細な説明

## 産業上の利用分野

本発明は、録音下、特に自動車の中、工場の中などでの音声認識用の音声区間検出方法に関する。

## 従来の技術

音声認識装置を実現する上で、音声区間の検出

は非常に重要であり、以後の認識結果に大きな影響を及ぼす。音声区間検出の一般的なものとしては、例えば「音声認識」(新見著、共立出版、p. 68~69)に示されるように2つの閾値を用いる方法がある。

## 発明が解決しようとする課題

ところが、一般に母音“い”は他の母音に比べて音声パワーが小さいので、従来法によると欠落が生じ、漏認識してしまうことがある。このような母音“い”の欠落対策として例えば特開昭60-260036号公報に示されるように、区間検出処理前に母音“い”を検出する手段を設け、母音“い”と判断した場合、区間検出のための閾値を下げるという方法があるが、母音“い”を判断する手段を設けなければならず、コストや処理速度の点で問題が残る。

## 課題を解決するための手段

新音を集音して電気信号に変換し、変換された

## 特開平 4-163497 (2)

信号を基に音声の存在する区間を検出するようにした音声区間検出方法において、音声パワーの小さい母音のホルマント周波数近傍の周波数成分を強調するプリアンファシスを通じた出力を用いて音声の存在する区間を検出するようにした。

作用

母音“い”を強調するプリアンファシスをかけた出力信号を用いて音声区間を検出するので、音声パワーの小さい母音“い”の欠落の可能性が低下し、音声区間の誤検出が少なくなる。

実施例

本発明の一実施例を図面に基づいて説明する。基本的には、音声入力部 1 と本実施例の特徴とするプリアンファシス 2 と閾値計算部 3 と区間検出部 4 と音声認識部 5 よりなる。

このような構成において、本実施例の音声区間検出処理は第 2 図に示すフローチャートに従い行われる。

アシス 2 の出力  $P_v$  とする。

$$Y(f) = w(f) \cdot X(f)$$

$$P_v = \sum_{f=1}^m Y(f)$$

なお、上式中の重み係数  $w(f)$  は母音“い”の第 1, 2 ホルマントなどを強調するもので、発声者が男性の場合は  $100 \sim 300 \text{ Hz}$  (第 1 ホルマント)、 $1.9 \sim 2.4 \text{ kHz}$  (第 2 ホルマント) 付近で大きな値となる。第 4 図は重み係数  $w(f)$  の例を示したものである。

また、このようなプリアンファシスの代りに、第 4 図に示すような特性を持つフィルタなどを用いてもよい。

第 5 図は、例えば“北見(きたみ)”と発声したときの入力信号  $x(t)$  の短時間パワーの時系列  $P_x(n)$  と、プリアンファシスの出力の時系列  $P_v(n)$  とを示している。なお、ここで用いるパワーとは、入力信号の絶対値、或いは目録値の短

まず、入力音の取込みを行う。音声入力部 1 はマイクロフォンのような音響・電気信号変換器によるもので、音声を集音して電気信号  $x(t)$  に変換する。

ついで、プリアンファシス計算処理を行う。プリアンファシス 2 は母音“い”などのように音声パワーの小さい音声をも大きくするために音声パワーの小さい母音のホルマント周波数近傍の周波数成分を強調するものであり、例えば第 3 図に示すように構成されている。このプリアンファシス 2 は帯域パワー検出部 6 で入力信号  $x(t)$  をバンドパスフィルタ群、或いは F E T などを用いて、 $x(1) \sim x(m)$  なる  $m$  個の帯域に分割した後、 $10 \text{ msec}$  程度の短時間スペクトル  $X(f)$  を検出し (ただし、 $f = 1, 2, \dots, m$ )、重み付けパワー検出部 7 で各帯域パワーを次式のように重み係数記憶部 8 に記憶しておいて重み係数  $w(f)$  を用いて、重み付けし、和をとったものをプリアンファ

時間平均値である。

ついで、閾値計算部 3 による閾値計算を選択的に行う。即ち、音声が存在しない区間におけるプリアンファシス 2 の出力信号  $P_v$  から閾値  $Th$  を計算し、記憶する。例えば、

$$Th = \alpha \cdot P_v + \beta$$

として求める。

一方、区間検出部 4 によりプリアンファシス 2 の出力信号  $P_v$  を用いて区間検出を行う。区間検出の方法としては、プリアンファシス 2 の出力信号  $P_v$  が閾値計算部 3 で求められた閾値  $Th$  を一定時間以上越えた区間を音声とする方法、その他の方法でよい。

この区間検出部 4 の出力信号と入力信号  $x(t)$  とは音声認識部 5 に入力され、入力信号  $x(t)$  の内で区間検出部 4 で区間検出された信号のみを用いて音声認識される。認識方法は、例えば「2 値の T S P を用いた単語音声認識システムの開発」

## 特開平4-163497(3)

(安田時朗他著、電気学会論文誌C108巻、昭和63年10月号、p858～865)に示されるような音声認識システムでよいが、その他の周知の音声認識システムでもよい。

## 発明の効果

本発明は、上述したように母音“い”のように音声パワーの小さい母音のホルマント周波数近傍の周波数成分を強調するプリアンファスを通した出力を用いて音声の存在する区間を検出するようにしたので、語頭や語尾に音声パワーの小さい母音“い”を含む単語を認識するに際しても、処理速度の遅れを伴うことなく、母音“い”の欠落の可能性を低下させて、音声区間の誤検出を減少させることができる。

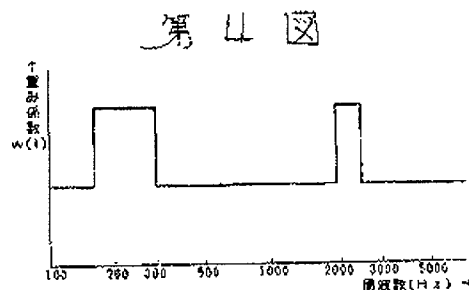
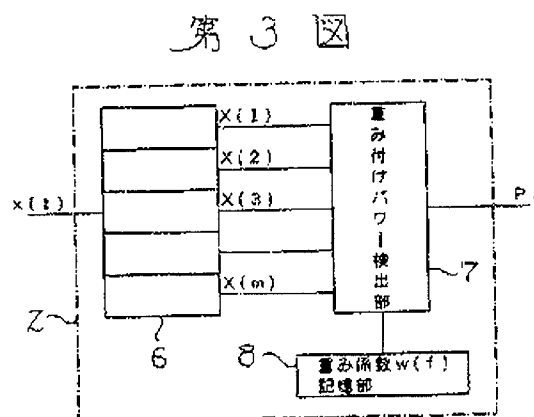
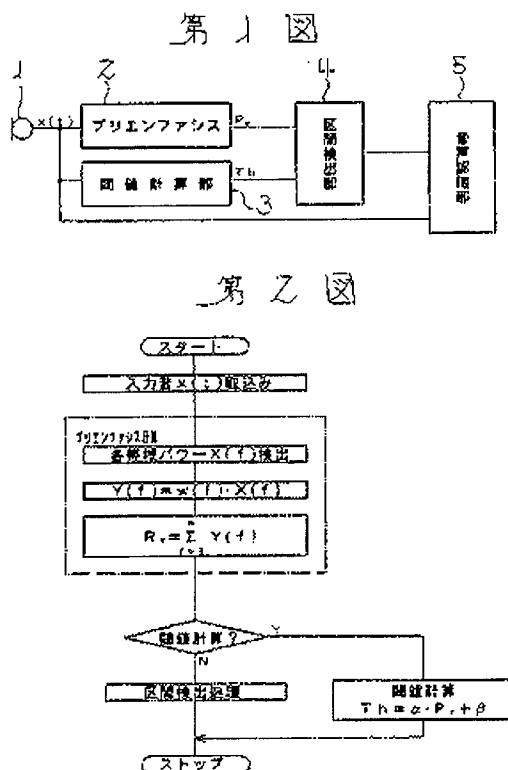
はプリアンファシスの構成を示すブロック図、第4図は重み係数例を示す特性図、第5図は具体的なプリアンファシス出力例を示す特性図である。

出願人 株式会社 リコー  
代理人 柏 木



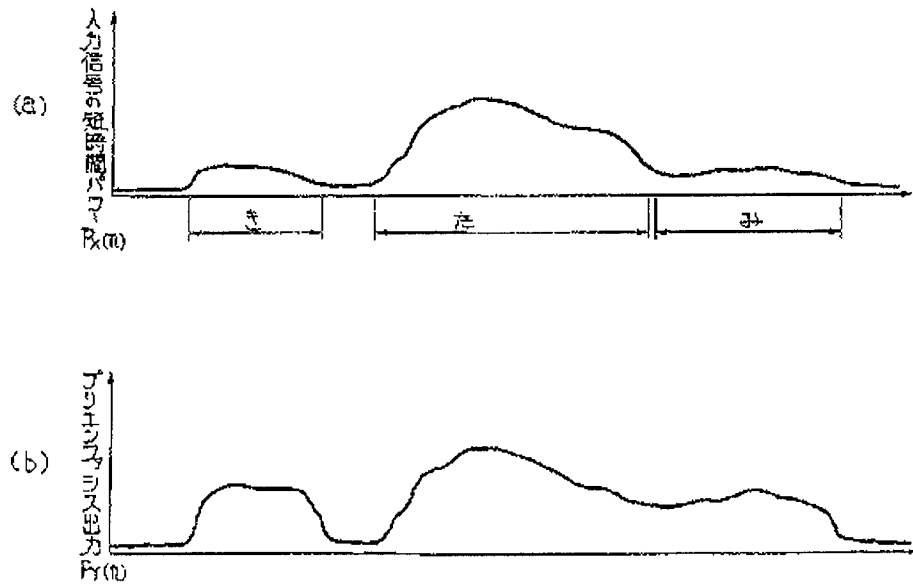
## 4. 図面の簡単な説明

図面は本発明の一実施例を示すもので、第1図はブロック図、第2図はフローチャート、第3図



特開平 4-163497 (4)

第 5 図



特開平4-163497

【公報種別】特許法第17条の2の規定による補正の掲載

【部門区分】第6部門第2区分

【発行日】平成11年(1999)1月22日

【公開番号】特開平4-163497

【公開日】平成4年(1992)6月9日

【年道号数】公開特許公報4-1635

【出願番号】特願平2-289113

【国際特許分類第6版】

G10L	3/00	513
	3/02	301
	9/00	
	9/02	301

【F I】

G10L	3/00	513 A
	3/02	301 Z
	9/00	D
	9/02	301 A

特許 0992 特開 005 特許 (第 6 版)

平成 9 年 6 月 6 日

特許庁長官 渡部 武光 様

F. 事件の表示

特許 0992 特開 005 特許 (第 6 版)

2. 特許する旨

事件との関係 特許出願人

住 所 東京都港区南青山5丁目9番15号 株式会社(新行山)

番 号 674 株式会社リコー

代表者 渡部 武光

3. 代理人

〒107

住 所 東京都港区南青山5丁目9番15号 株式会社(新行山)

番 号 03(3408)4535

氏 名 7211 株式会社 鈴木 明子

4. 訂正命令の交付

なし

5. 訂正の対象

明細書、図面

6. 訂正の内容

別紙のとおり

この訂正に關し、特許官及び特許官の意見を下記のように記載する。

訂

1. 明細書中、第4頁(第7行目の「小さい音声を」を「小さい音声を」に訂正する。

2. 明細書中、第4頁(第13行目の「1/1000」を「1/1000」に訂正する。

3. 特許書中、第8頁(第10行目の「ブレンファシス曲力図を示す」を「ブレンファシスの図を示す」に訂正する。

4. 図面中、第1図を別紙のように訂正する。

特開平4-163497

